

---

# **TECHNICKÁ UNIVERZITA V LIBERCI**

Fakulta mechatroniky, informatiky a mezioborových studií

Studijní program: B2612 – Elektrotechnika a informatika

Studijní obor: 1802R022 – Informatika a logistika

## **Slepá separace akustických zdrojů pomocí lokalizace**

## **Blind separation of acoustical sources by means of localization**

### **Bakalářská práce**

Autor: **Martin Češka**

Vedoucí práce: Ing. Zbyněk Koldovský, Ph. D.

V Liberci 29. 5. 2009

## Prohlášení

Byl jsem seznámen s tím, že na mou bakalářskou práci se plně vztahuje zákon č. 121/2000 o právu autorském, zejména § 60 (školní dílo).

Beru na vědomí, že TUL má právo na uzavření licenční smlouvy o užití mé bakalářské práce a prohlašuji, že **s o u h l a s í m** s případným užitím mé bakalářské práce (prodej, zapůjčení apod.).

Jsem si vědom toho, že užít své bakalářské práce či poskytnout licenci k jejímu využití mohu jen se souhlasem TUL, která má právo ode mne požadovat přiměřený příspěvek na úhradu nákladů, vynaložených univerzitou na vytvoření díla (až do jejich skutečné výše).

Bakalářskou práci jsem vypracoval samostatně s použitím uvedené literatury a na základě konzultací s vedoucím bakalářské práce a konzultantem.

Datum : 29. 5. 2009

Podpis:

## **Poděkování**

Chtěl bych touto cestou poděkovat svému vedoucímu Ing. Zbyňkovi Koldovskému, Ph. D., díky kterému tato práce mohla vzniknout, za všechny cenné rady a doporučení, které mi dával v průběhu zpracovávání bakalářské práce, a za jeho ochotu a vstřícný přístup při poskytování informací.

## **Abstrakt**

Bakalářská práce je zaměřena na lokalizaci komponent akustických zdrojů pomocí metody slepé separace zdrojů. Cílem této metody je prostřednictvím separačního procesu obnovit původní zdrojové signály z jejich směsí. Úkolem práce je na základě teoretických znalostí a s využitím vývojového prostředí MATLAB navrhnout způsob, jakým by mohla být tato metoda využita k odhadnutí směru, ze kterého přichází jednotlivé komponenty akustických zdrojů k mikrofonom. Druhým úkolem je nalezení způsobu využití odhadu směru komponent, pomocí něhož by mohla být metoda slepé separace zdrojů vylepšena.

**Klíčová slova:** slepá separace, akustický zdroj, směs signálů, mikrofon.

## **Abstract**

The bachelor thesis is focused on localization of components of acoustical sources using blind source separation method. Objective of this method is to retrieve original source signals from their mixture by means of separation procedure. The goal of the thesis is to propose a solution, based on a theoretical knowledge and a usage of MATLAB developing environment, in which this method could be used to estimate a direction from which particular components of acoustical sources arrive at microphones. The second goal is to enhance blind source separation method by the estimation of components arrival direction.

**Keywords:** blind separation, acoustical source, mixture of signals, microphone.

# Obsah

Prohlášení.....	3
Poděkování.....	4
Abstrakt.....	5
Abstract.....	6
Seznam obrázků a grafů.....	9
Seznam symbolů, zkratk a termínů .....	10
1 Úvod.....	11
2 Metoda slepé separace zdrojů .....	12
2.1 Cocktail-party problém .....	12
2.1.1 Matematický popis cocktail-party problému.....	14
2.1.2 Analýza nezávislých komponent (ICA) .....	15
2.1.3 Analýza hlavních komponent (PCA).....	16
2.1.4 Výpočetní analýza zvukové scény (CASA) .....	17
2.1.5 Praktické využití řešení cocktail-party problému.....	17
2.2 Hlavní kroky výchozí separační metody.....	18
2.2.1 Rozklad pomocí ICA.....	20
2.2.2 Shlukování nezávislých komponent .....	21
2.2.3 Rekonstrukce .....	22
3 Praktická aplikace metody BSS při odhadu směru signálů a její možné vylepšení ....	24
3.1 Myšlenka řešení stanovených cílů .....	24
3.1.1 Původní provedení metody BSS.....	26
3.1.2 Využití metody BSS k odhadu směru komponent .....	27
3.1.3 Vylepšení metody BSS s využitím odhadu směru zdrojů .....	29
3.2 Implementace řešení v prostředí MATLAB .....	31

3.2.1 Vývojové prostředí MATLAB .....	31
3.2.2 Vzhled aplikace a zdrojový kód .....	33
4 Experimenty.....	38
4.1 Vlastnosti a kritéria provedených experimentů.....	38
4.2 Vyhodnocení provedených separací .....	39
5 Závěr .....	43
Seznam použité literatury .....	44

## Seznam obrázků a grafů

Obrázek 1 Blokový diagram znázorňující cocktail-party problem.....	13
Obrázek 2 MEG – magnetoencefalografie.....	18
Obrázek 3 Schematické znázornění procedury separace.....	19
Obrázek 4 Šíření zvukových signálů od zdrojů k mikrofonom.....	25
Obrázek 5 Příklad matice znázorňující energii komponent na mikrofonech.....	28
Obrázek 6 Příklad matice znázorňující směrovost komponent.....	30
Obrázek 7 Rozhraní aplikace MATLAB.....	31
Obrázek 8 Dialogové okno aplikace pro BSS z prostředí MATLAB.....	33
Obrázek 9 Řádky 1 – 17 zdrojového kódu algoritmu v prostředí MATLAB.....	34
Obrázek 10 Řádky 18 – 42 zdrojového kódu algoritmu v prostředí MATLAB.....	34
Obrázek 11 Řádky 43 – 54 zdrojového kódu algoritmu v prostředí MATLAB.....	35
Obrázek 12 Řádky 55 – 71 zdrojového kódu algoritmu v prostředí MATLAB.....	36
Obrázek 13 Řádky 72 – 88 zdrojového kódu algoritmu v prostředí MATLAB.....	36
Obrázek 14 Řádky 89 – 100 zdrojového kódu algoritmu v prostředí MATLAB.....	37
Obrázek 15 Pole mikrofonom pro pořízení smíšených audio záznamů.....	39
Graf 1 Hodnoty SIR a SDR v závislosti na použité dvojici mikrofonom.....	40
Graf 2 Hodnoty SIR a SDR v závislosti na použitém nastavení.....	41
Graf 3 Hodnoty SIR a SDR v závislosti na zkrácení délky separovaného signálu.....	42



## Seznam symbolů, zkratek a termínů

- AJD – Approximate Joint Diagonalization (přibližná společná diagonalizace)
- BGL – Block Gaussian Likelihood (bloková Gaussova pravděpodobnost)
- BSS – Blind Source Separation (slepá separace zdrojů)
- CASA – Computational Auditory Scene Analysis (výpočetní analýza zvukové scény)
- EFICA – Efficient Fast Independent Component Analysis (účinná rychlá analýza nezávislých komponent)
- FastICA – Fast Independent Component Analysis (rychlá analýza nezávislých komponent)
- FFT – Fast Fourier Transform (rychlá Fourierova transformace)
- FIR – Final Impulse Response (konečná odezva impulsu)
- ICA – Independent Component Analysis (analýza nezávislých komponent)
- Infomax – Information Maximization (informační maximalizace)
- JADE – Joint Approximate Diagonalization of Eigen-matrices (společná přibližná diagonalizace charakteristických matic)
- KLТ – Karhunen-Loève Transform (Karhunen-Loèevova transformace)
- MATLAB – MATrix LABoratory (laboratoř s maticemi)
- MEG – Magnetoencephalography (magnetoencefalografie)
- MISO – Multiple Input Single Output (vícenásobný vstup, jeden výstup)
- PCA – Principal Component Analysis (analýza hlavních komponent)
- POD – Proper Orthogonal Decomposition (náležitá ortogonální dekompozice)
- SDR – Signal-To-Distortion ratio (poměr signálu ke zkreslení)
- SIR – Signal-To-Interference ratio (poměr signálu k interferenci)
- SOBI – Second-Order Blind Identification (slepá identifikace druhého řádu)
- WASOBI – Weight-Adjusted Second-Order Blind Identification (slepá identifikace druhého řádu s upravenými váhami)

# 1 Úvod

Tématem bakalářské práce je separace na základě lokalizace akustického zdroje pomocí metody „Blind Source Separation“ neboli slepé separace zdrojů. Pojem slepá chápeme tak, že metoda pracuje bez jakékoli znalosti akustiky prostředí, polohy zdrojů a mikrofonů či informací o původních separovaných signálech. Jedná se v podstatě o skupinu metod číslicového zpracování signálů, jejichž cílem je pomocí separačního procesu obnovit původní zdrojové signály z jejich směsí.

Pro lepší pochopení je možné si představit například situaci, kdy se v jedné místnosti nachází několik současně mluvících osob a snahou je zaměřit se pouze na jednu z probíhajících diskuzí – tato situace se v teorii označuje jako cocktail-party problém. Pomocí metody slepé separace zdrojů pak dokážeme ze záznamu hovoru oddělit řeč jednotlivých lidí. V praxi tato metoda nachází široké uplatnění zejména v oblasti telekomunikací, biomedicíny, zpracování řeči a jiných audio signálů, zpracování geoseizmických dat a podobně.

Úkolem práce je navrhnout způsob, jakým by mohla být tato metoda využita ke zjištění směrové charakteristiky jednotlivých komponent akustických zdrojů a následně nalézt možné využití odhadu směru těchto komponent, pomocí něhož by mohla být metoda pro slepou separaci vylepšena. Navržené řešení uvažuje dva mikrofony a předpokládá právě dva zdroje.

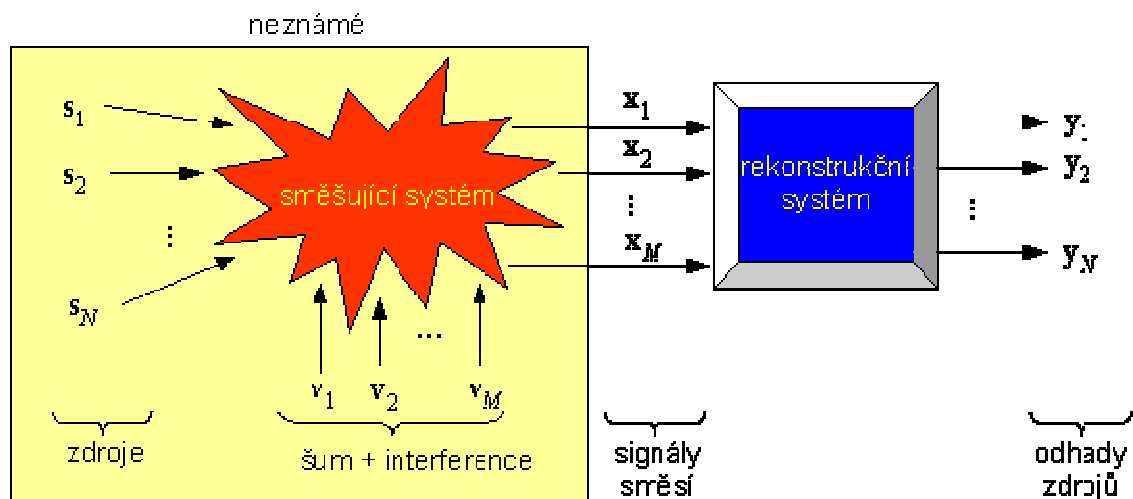
Na základě stanovených cílů je práce strukturována do několika základních částí. V části teoretické jsou obsaženy všechny podstatné informace o metodě slepé separace zdrojů. Následující kapitola popisuje myšlenku vlastního řešení úkolů, tedy jak využít metodu slepé separace k odhadnutí směru komponent a za pomoci tohoto odhadu vylepšit samotnou metodu slepé separace zdrojů. Dále je vysvětlena metoda, ze které vychází řešení úkolu, následuje popis vlastního řešení a nakonec jeho implementace v prostředí programu MATLAB. Závěrečná část práce se věnuje experimentům s audio nahrávkami a zhodnocení dosažených výsledků.

## 2 Metoda slepé separace zdrojů

Následující kapitola je zaměřena na metodu slepé separace zdrojů. V první části je popsán jeden ze základních problémů (tak zvaný „cocktail-party problém“), který se uvedenou metodou řeší, včetně jeho matematického vysvětlení; dále tři základní metody, kterými je možné cocktail-party problém řešit, a nakonec praktické využití uvedených technologií. V druhé části kapitoly je vysvětlena samotná procedura výchozí separační metody, která pracuje v časové oblasti a je představena v [3]. Skládá se z analýzy zaznamenaných akustických signálů, respektive jejich komponent, dále z jejich následného shlukování na základě vzájemné podobnosti a nakonec z rekonstrukce odezev původních zdrojů na mikrofonech. Tato metoda byla použita jako základ a navržené řešení problému z ní vychází. Existují také další metody, například fungující ve frekvenční oblasti.

### *2.1 Cocktail-party problém*

Problém separace zdrojů v oblasti zpracování digitálního signálu spočívá v situaci, kdy je několik signálů smícháno dohromady, a úkolem je zjistit, jak vypadaly původní signály. Klasickým příkladem je cocktail-party problém, kdy spolu v místnosti hovoří skupina lidí a někdo se snaží poslouchat pouze jednu z probíhajících diskuzí. Lidský mozek bez obtíží zvládne tento typ sluchové separace zdroje, ale v oblasti zpracování digitálního signálu se jedná o velmi složitý problém. [13]



Obrázek 1 Blokový diagram znázorňující cocktail-party problém

Zdroj: EKSLEK, V. *Analýza hlavních komponent v problematice separace naslepo* [online]. [cit. 16.5.2009]. URL: <<http://www.elektrorevue.cz/clanky/05029/index.html>>.

Jako první analyzoval cocktail-party problém Edward Colin Cherry (1914 – 1979). Hlavní přínos tohoto britského vědce spočíval v zaměření na sluchovou pozornost, konkrétně týkající se cocktail-party problému.

Jeho analýza se zabývala problémem sledování pouze jedné konverzace, zatímco mnoho jiných konverzací probíhá okolo. Cherry ke studování tohoto problému používal stínové úlohy, kdy účastníkovi pokusu přehrával do levého a pravého ucha dvě odlišné zvukové zprávy a instruoval ho, aby se věnoval pouze jedné z nich. Účastník pokusu dostal následně za úkol zopakovat (stínovat) tu zvukovou zprávu, které věnoval pozornost.

Cherry zjistil, že účastníci jeho pokusu získali velmi málo informací ze zprávy, které nevěnovali pozornost. Byly zjištěny hmotné charakteristiky, ale sémantické (významové) nikoli. Cherry následně usoudil, že zvukové informace, které není věnována pozornost, se dostane velmi nepatrného zpracování a že lidé používají hmotné odlišnosti mezi zprávami k výběru, kterou zprávou se zabývat. [8]

### 2.1.1 Matematický popis cocktail-party problému

Vzhledem k šíření zvuku v přirozeném akustickém prostředí je směšovací proces nastávající u cocktail-party problému popsán vztahem

$$x_i(n) = \sum_{j=1}^d \sum_{\tau=0}^{M_{ij}-1} h_{ij}(\tau) s_j(n - \tau), \quad i = 1, \dots, m, \quad (1)$$

kde  $d$  je počet nezávislých zvukových zdrojů,  $m$  počet mikrofónů,  $x_1(n), \dots, x_m(n)$  sledované signály na mikrofonech,  $s_1(n), \dots, s_d(n)$  neznámé původní (audio) signály a  $h_{ij}$  impulzní odezvy zdroj-senzor, každá o délce  $M_{ij}$ .

Získat původní signály bez jakékoli předchozí znalosti je všeobecně nemožné. Nicméně za použití předpokladu vzájemné nezávislosti původních zdrojů, čemuž obvykle dobře odpovídají vlastnosti promluv řečníků, je možné použít metodu slepé separace. Tedy separace může být provedena metodou analýzy nezávislých komponent (ICA).

Akustické zdroje mají zpravidla neznámou dočasnou strukturu, což způsobuje neurčitosti v řešení jejich slepé separace. Hlavním cílem je tedy spíše odhadnout odezvy zdrojů na mikrofonech než odhadnutí zdrojů samotných. Pro  $k$ -tý zdroj a  $i$ -tý mikrofón je odezva

$$s_i^k(n) = \sum_{\tau=0}^{M_{ij}-1} h_{ij}(\tau) s_k(n - \tau), \quad (2)$$

a je odhadována jako výstup MISO (Multiple Input Single Output) filtru o délce  $L$ .

$$\hat{s}_i^k(n) = \sum_{j=1}^m \sum_{\tau=0}^{L-1} w_{ij}^k(\tau) x_j(n - \tau). \quad (3)$$

Pro řešení cocktail-party problému bylo navrženo několik metod, ale intenzivní výzkum v této oblasti nadále pokračuje. Z úspěšných postupů můžeme jmenovat například analýzu nezávislých komponent (ICA), která byla použita při řešení této práce, analýzu hlavních komponent (PCA) a výpočetní analýzu zvukové scény (CASA), která se snaží dosáhnout zvukové separace zdrojů pomocí metody založené na lidském sluchu. [3, 13]

### **2.1.2 Analýza nezávislých komponent (ICA)**

Analýza nezávislých komponent jako jedna z možností řešení cocktail-party problému je statistická a výpočetní metoda pro odhalování skrytých faktorů, které jsou základem skupin náhodných proměnných, měření nebo signálů.

ICA definuje model pro pozorování mnoha náhodně proměnných dat, který je typicky stanoven jako velká databáze vzorků. V tomto modelu jsou datové proměnné považovány za lineární směsi nějakých neznámých skrytých proměnných a není znám ani systém směřování. Skryté proměnné jsou považovány za ne-Gaussovské a vzájemně nezávislé a jsou nazývány nezávislé komponenty pozorovaných dat. Tyto nezávislé komponenty, také nazývané zdroje nebo faktory, mohou být nalezeny metodou ICA.

Analýza nezávislých komponent je povrchně spřízněná s analýzou hlavních komponent a faktorovou analýzou. Nicméně ICA je mnohem silnější metoda schopná nalézt základní faktory nebo zdroje, i když tyto ostatní metody kompletně selžou.

Data analyzovaná metodou ICA mohou pocházet z mnoha odlišných oblastí praktického využití, zahrnující například digitální obrazy, databáze dokumentů, ekonomické ukazatele a psychometrická měření. V mnoha případech jsou hodnoty zadány jako skupiny paralelních signálů nebo časových oddílů. Tento problém charakterizujeme termínem slepá separace zdrojů (BSS). Typickými příklady jsou směsi současně probíhajících řečových signálů zaznamenaných několika mikrofony, mozkové

vlny nahrané mnoha senzory, rušivé rádio signály přicházející na mobilní telefon nebo paralelní časové řady získané z nějakého průmyslového postupu.

Další klasický příklad je právě cocktail-party problém, kdy jsou výchozí řečové signály separovány ze vzorku dat, který obsahuje skupinu současně hovořících lidí v místnosti. V teorii je problém zjednodušen předpokladem neexistence ozvěny a časového zpoždění, v praxi je však potřeba tyto faktory uvažovat. Důležitým faktem je, že pokud je přítomno  $N$  zdrojů, potřebujeme nejméně  $N$  mikrofónů k získání původních signálů. [2, 10]

### 2.1.3 Analýza hlavních komponent (PCA)

Analýza hlavních komponent je další z možností, jak řešit cocktail-party problém. Obsahuje matematickou proceduru, která transformuje určité množství pravděpodobně souvisejících proměnných na menší množství nesouvisejících proměnných zvaných hlavní komponenty.

První hlavní komponenta odpovídá co největší možné varianci dat a každá následující komponenta odpovídá co největší zbývající varianci dat. Cílem analýzy hlavních komponent je získat relativně malý počet vzájemně nezávislých lineárních kombinací (hlavních komponent), které nesou co nejvíce informací původních proměnných. Jinými slovy, hlavním účelem PCA je komprese dat a výběr signálů nesoucích nejvíce energie.

Hlavní komponenty PCA jsou náhodné proměnné s maximálními odlišnostmi sestavené z lineárních směrů vstupních prvků. Jelikož je PCA založena na statistikách druhého řádu, redukuje pouze statistickou závislost druhého řádu, což je ale pro separaci jednotlivých zdrojových signálů nedostačující. K úspěšné separaci je nezbytně nutné redukovat také statistické závislosti vyšších řádů, což umožňuje metoda ICA.

PCA metoda byla vynalezena v roce 1901 Karlem Pearsonem. V současnosti je používána převážně jako nástroj pro výzkumnou datovou analýzu a pro vytváření predikčních modelů. V závislosti na oboru aplikace je tato metoda také nazývána

diskrétní Karhunen–Loèvova transformace (KLT), Hotellingova transformace nebo POD (Proper Orthogonal Decomposition). [12]

#### **2.1.4 Výpočetní analýza zvukové scény (CASA)**

Výpočetní analýza zvukové scény je jednoduše řečeno studie analýzy zvukové scény za pomoci výpočetní techniky. CASA systémy jsou v podstatě odposlouchávací zařízení, která mají za cíl separovat směsi zvukových zdrojů stejnou cestou, jako to dělá člověk.

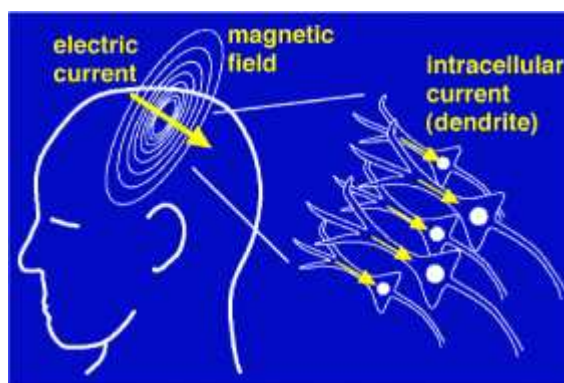
CASA se liší od systémů BSS v tom, že je (alespoň do určité míry) založena na mechanizmech lidského sluchového systému a používá tedy nejvýše dvě mikrofónové nahrávky z akustického prostředí. [9]

#### **2.1.5 Praktické využití řešení cocktail-party problému**

Výše uvedené technologie mají potencionální využití v oblastech automatického rozpoznání řeči a mluvčího v hlučném prostředí, pomůcek pro sluchově postižené a automatického přepisu hudebních audio nahrávek. V případě separace hudebních signálů je pro stereo mix relativně jednoduchých signálů dnes možné provést velmi přesnou separaci, ačkoli některé elementy přeci jen v signálu zůstanou.

Další z praktických využití, které je vyvíjeno v této oblasti, je měření magnetické aktivity mozku pomocí metody MEG (magnetoencefalografie) používané v medicíně. Tento druh zobrazování zahrnuje pečlivé měření magnetického pole okolo hlavy, které poskytne přesný prostorový obraz vnitřku hlavy. Ovšem externí zdroje elektromagnetického pole jako například hodinky na paži vyšetřované osoby velmi významně snižují přesnost měření. Použitím technik separace zdrojů na měřené signály je možné odstranit nežádoucí elementy ze signálu. [9, 13]





Obrázek 2 MEG – magnetoencefalografie

Zdroj: WIKIPEDIA. *Magnetoencephalography* [online]. [cit. 16.5.2009].

URL: <<http://en.wikipedia.org/wiki/Magnetoencephalography>>.

## 2.2 Hlavní kroky výchozí separační metody

Níže uvedená metoda, která slouží jako podklad pro vlastní řešení této práce, vychází z modelu separace v časové oblasti. Společným znakem separačních metod v časové oblasti je to, že pracují s maticí časově posunutých signálů, naměřených na daném mikrofonním poli. Počet řádků v této matici je stanoven součinem počtu mikrofónů a počtu posunutí, který reprezentuje délku separujícího inverzního filtru. Tato metoda vykazuje velmi dobré výsledky i v případě použití relativně malého počtu vzorků signálů z mikrofónů a krátkého separujícího inverzního filtru.

Metoda pracuje následujícím způsobem:

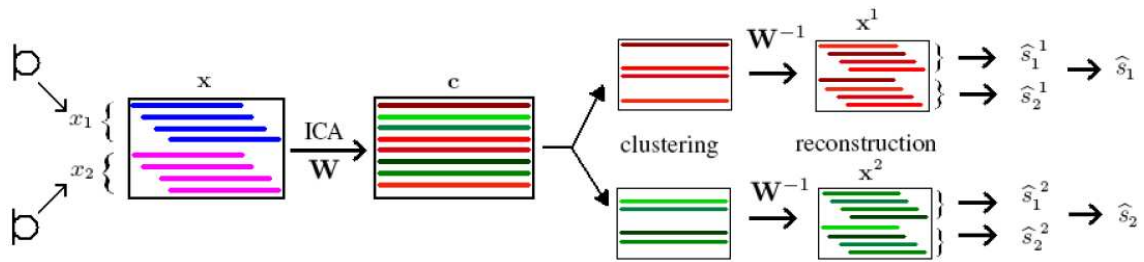
1. ICA transformace (provedená metodou ICA) je aplikována na zvolenou dimenzi podprostoru signálů určenou součinem  $m \cdot L$ , která se skládá z těchto časově posunutých vzorků signálů:

$$\mathbf{x}(n) = [x_1(n), x_1(n-1), \dots, x_1(n-L+1), x_2(n), \dots, x_m(n-L+1)]^T. \quad (4)$$

Výsledné (nezávislé) komponenty mohou být chápány jako výstupy MISO FIR (Final Impulse Response) filtrů  $m \cdot L$  o délce  $L$ , které jsou vybrány tak, aby byly výstupy co nejvíce vzájemně nezávislé.

2. Komponenty jsou seskupeny za použití algoritmu shlukování tak, aby komponenty sdílející společný shluk byly (v ideálním případě) filtrované verze jedna druhé a patřily ke stejnému původnímu akustickému zdroji.
3. K získání individuálních odezev zdrojů je na shluky jednotlivých komponent použita procedura rekonstrukce.

Níže znázorněné schéma ilustruje, jakým způsobem separační procedura zpracovává skupinu signálů ze dvou mikrofónů a na konci ji rozděluje na dva samostatné zdroje. Jedná se o případ, kdy je v rekonstrukci použito binární váhování (blíže vysvětleno v kapitole 2.2.3).



Obrázek 3 Schematické znázornění procedury separace

Zdroj: KOLDOVSKÝ, Z., TICHAVSKÝ, P. *Time-domain Blind Audio Source Separation Using Advanced Component Clustering and Reconstruction*. Proc. of The Joint Workshop on Hands-free Speech Communication and Microphone Arrays (HSCMA 2008), pp. 216 – 219. Trento, Italy, 2008. ISBN 978-1-4244-2338-5.

Následující tři podkapitoly jsou věnovány detailnímu popisu výše uvedených kroků separační procedury. [3, 6]

### 2.2.1 Rozklad pomocí ICA

Od komponent získaných prostřednictvím ICA se všeobecně očekává, že jejich vzájemné časoprostorové působení bude co nejvíce vyrušeno. V ideálním případě se jedná o různě filtrované kopie nezávislých původních zdrojů  $s_1(n), \dots, s_d(n)$ . K realizaci rozkladu ICA lze použít jednu ze tří základních tříd ICA algoritmů.

První třída, reprezentovaná algoritmy JADE (Joint Approximate Diagonalization of Eigen-matrices), FastICA, Infomax (Information Maximization) a nově EFICA (Efficient Fast ICA), je založena na ne-Gaussovskosti původních zdrojů. Druhá třída, reprezentovaná algoritmy SOBI (Second-Order Blind Identification) a WASOBI (Weight-Adjusted SOBI), spoléhá na výrazné zabarvení zdrojů (spektrální diversitu), a třetí třída, reprezentovaná BGL (Block Gaussian Likelihood), je založena na jejich nestacionaritě. Zatímco první třída používá statistiku dat vyššího řádu (nelineární transformace), druhé dvě třídy jsou založeny na statistice druhého řádu. BGL spočívá v rozdělení získaných signálů na určitý počet nepřesahujících segmentů, výpočtu kovariančních matic signálů pro každý segment a přibližné společné diagonalizaci (AJD) těchto matic.

Algoritmus EFICA je v literatuře popisován jako velmi dobrá ICA transformace vztahu (4). Je ovšem možné využít i implementaci rychlého BGL algoritmu, který používá nový AJD algoritmus. Tato implementace umožňuje ICA transformaci o délce dat 6000 vzorků s  $m = 2$ ,  $L = 20$  a při uvažovaných 20-ti blocích přibližně za jednu sekundu, což odpovídá AJD s 20-ti maticemi o velikosti 40x40. EFICA je pomalejší a umí provést podobnou separaci za 6 sekund. [3]

## 2.2.2 Shlukování nezávislých komponent

Nechť  $x$  označuje datovou matici  $x(n)$  pro  $n = 1, \dots, N$ . Jakmile je ICA aplikována na  $x$ , což dává demixující transformaci  $W$ , získáme  $m \cdot L$  libovolně seřazených komponent v  $c = Wx$ . Bez jakékoli ztráty všeobecné platnosti znormalizujeme řádky  $W$ , abychom dosáhli toho, že všechny komponenty budou mít stejnou jednotkovou energii.

Klíčovým předpokladem této metody je, že každá komponenta je filtrovanou verzí jednoho z  $d$  původních zdrojů, tudíž mohou být komponenty seskupovány do  $d$  shluků tak, aby každý z nich odpovídal jednomu původnímu zdroji. Proto je vhodné změřit podobnost mezi komponentami, které odrážejí, jak může být filtrovaná verze jedné z komponent blízko jiné komponentě a naopak.

Nechť  $c_j(n)$  označuje  $j$ -tou komponentu,  $j = 1, \dots, m \cdot L$ . Měření podobnosti mezi  $i$ -tou a  $j$ -tou komponentou je definováno [3] jako:

$$D_{ij} = \hat{E}[\mathbf{P}_i c_j(n)]^2 + \hat{E}[\mathbf{P}_j c_i(n)]^2, \quad (5)$$

kde  $P_i$  označuje projektor na podprostor generovaný posunutími  $i$ -té komponenty:

$$c_i(n - L + 1), \dots, c_i(n + L - 1), \quad (6)$$

a  $\hat{E}$  označuje operátor aritmetického průměru přes dostupné vzorky signálu. Operátor projekce je dán jako

$$\mathbf{P}_i = \mathbf{I} - \mathbf{C}_i(\mathbf{C}_i^T \mathbf{C}_i)^{-1} \mathbf{C}_i^T \quad (7)$$

kde  $C_i$  se skládá ze zpožděných verzí  $c_i$  jako ve vztahu (6). Matice  $C_i^T C_i$  je přibližně rovna násobku autokovarianční matice z  $c_i$  o velikosti  $(2L - 1) \times (2L - 1)$ . Tuto matici je možno spočítat prostřednictvím FFT (Fast Fourier Transform) a invertovat pomocí

Levinsonova algoritmu, obojí velice rychle. Tímto způsobem může být například výpočet všech vzdáleností  $D_{ij}$ , kdy  $N = 6000$ ,  $m = 2$  a  $L = 20$ , proveden přibližně za 0,5 sekundy.

Matice složená z prvků  $D_{ij}$  je poté použita jako vstup pro standardní algoritmus aglomeračního hierarchického shlukování. Shlukování začíná s  $m \cdot L$  množinami o jednom prvku. V každém kroku je počet shluků snižován o jeden, a to sloučením dvou shluků s nejmenší vzájemnou vzdáleností, která je vypočítána jako průměrná vzdálenost mezi jednotlivými komponentami ve shluku. Z toho plyne, že počet původních shluků je postupně snižován, dokud se nerovná vypočtenému (případně odhadnutému) počtu zdrojů  $d$ . [3]

### 2.2.3 Rekonstrukce

Rekonstrukce má za cíl transformovat komponenty všech shluků do odezev. Předpokládá se, že jakákoli komponenta (v jakémkoli shluku) může přispět k rekonstrukci jakéhokoli zdroje:

$$\mathbf{x}^i = \mathbf{W}^{-1} \text{diag}[\lambda_1^i, \dots, \lambda_{mL}^i] \mathbf{c} \quad (8)$$

kde  $\lambda_l^i$  je váha pro rekonstrukci  $i$ -tého zdroje za použití  $l$ -té komponenty,  $i = 1, \dots, d$  a  $l = 1, \dots, mL$ . Na závěr, jelikož struktura  $\mathbf{x}^i$  by měla být podobná struktuře  $\mathbf{x}$ , je možné získat požadované odezvy z  $\mathbf{x}^i$  jako:

$$\hat{s}_k^i(n) = \sum_{p=1}^L \mathbf{x}_{(k-1)L+p}^i(n+p-1). \quad (9)$$

Binární váhování komponent je velice jednoduché:  $\lambda_l^i$  je stanovena jako jednička nebo nula podle toho, zda  $l$ -tá komponenta náleží do  $i$ -tého shluku nebo ne. Tento postup nazýváme tvrdé (binární) váhování.

Binární váhování má jednu nevýhodu. Může se stát, že jedna nebo více komponent získaných v kroku 1 nebude jasně přiřazena ani k jednomu ze shluků získaných v kroku 2. Velmi často se to stává v případě nízko-frekvenčních komponent, protože je obtížné je oddělit, když jsou mikrofony blízko u sebe. Nicméně, tyto komponenty jsou přiřazeny k některému ze shluků navzdory tomu, že obsahují významný podíl energie z ostatních působících zdrojů.

Proto se doporučuje použít dokonalejší rekonstrukci za použití takzvaného „fuzzy“ váhování. To umožňuje všem komponentám přispět k jakémukoli z rekonstruovaných  $x^i$  definováním  $\lambda_l^i$  tak, že jejich hodnota odráží vztah  $l$ -té komponenty k  $i$ -tému shluku. Pro tento účel lze definici váhy  $\lambda_l^i$  vyjádřit jako:

$$\lambda_l^i = \left( \frac{\sum_{j \in K_i, j \neq l} D_{lj}}{\sum_{j \notin K_i, j \neq l} D_{lj}} \right)^\alpha, \quad (10)$$

kde  $K_i$  zahrnuje ukazatele komponent v  $i$ -tém shluku, a  $\alpha$  je nastavitelný pozitivní parametr, který řídí „tvrdost“ váhování.

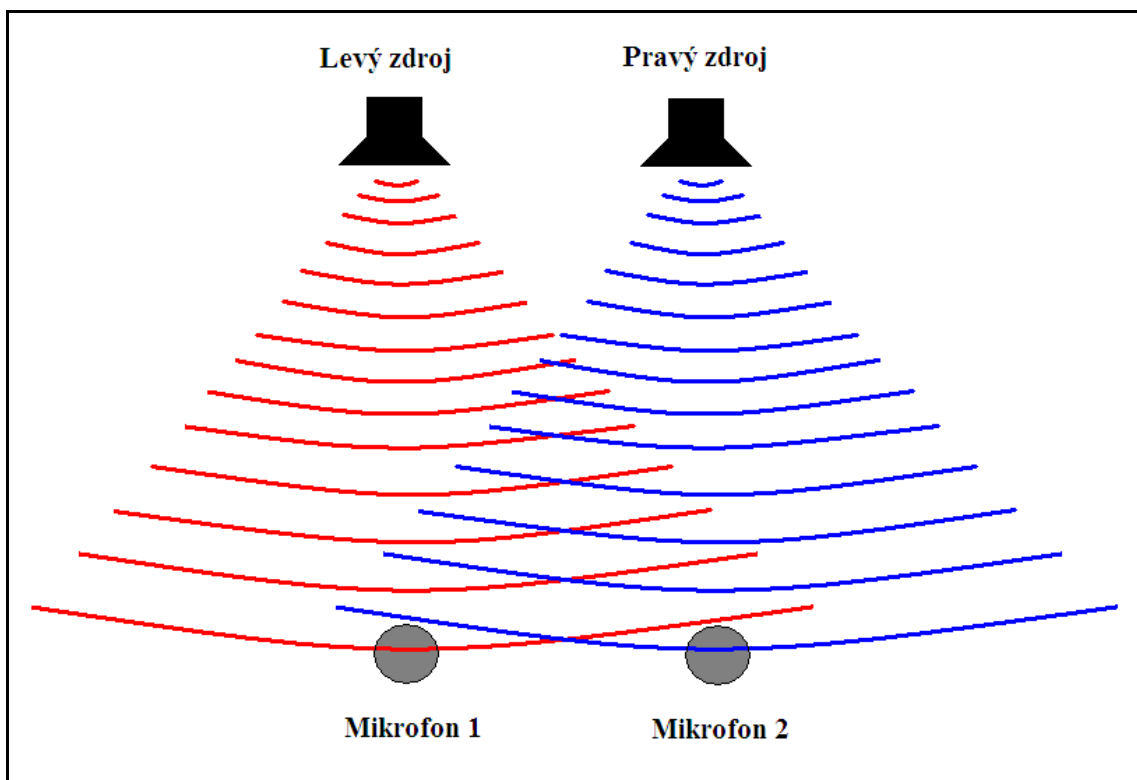
Stojí za povšimnutí, že v porovnání s kroky 1 – 2 už je rekonstrukce výpočetně nenáročná. [3]

## **3 Praktická aplikace metody BSS při odhadu směru signálů a její možné vylepšení**

Třetí kapitola, věnovaná praktickému řešení stanovených cílů, představuje stěžejní část bakalářské práce. Vyjadřuje základní myšlenku řešení úkolů, tedy jakým způsobem lze využít metodu slepé separace zdrojů při odhadu směru komponent a jak je možné za pomoci tohoto odhadu metodu slepé separace vylepšit. Následuje popis výchozí separační metody pro možnost srovnání jejího principu s metodou navrženou, která je vysvětlena hned poté. Další část kapitoly je zaměřena na představení interaktivního vývojového prostředí MATLAB, ve kterém byly zpracovány jednotlivé kroky separačního procesu a na závěr je vyobrazena a popsána vlastní implementace navrženého řešení v prostředí programu MATLAB.

### ***3.1 Myšlenka řešení stanovených cílů***

Myšlenka řešení vychází ze skutečnosti, že každý signál přichází k mikrofonu z jiného směru a tím pádem i z jiné vzdálenosti. To má za následek, že se také liší hlasitost, s jakou je ten který signál zaznamenán na konkrétním mikrofonu. Tuto hlasitost lze vyjádřit jako energii, se kterou signál dorazil k mikrofonu. Pokud je tedy zaznamenán smíšený záznam dvou signálů ze dvou mikrofonů (předpoklad dvou mikrofonů a právě dvou zdrojů je nutný pro funkčnost navržené metody), je možné po vhodně provedené separaci a označení jednoho z mikrofonů jako levý a druhého jako pravý určit, který signál přišel z levého zdroje a který ze zdroje pravého.



Obrázek 4 Šíření zvukových signálů od zdrojů k mikrofónům

Zdroj: vlastní zpracování

Každý signál se skládá z určitého počtu komponent. Pro jednotlivé komponenty platí stejná pravidla jako pro celý signál, tedy že každá komponenta má vlastní směr a energii na mikrofónu. Pokud je k dispozici jejich směs, je možné za použití vhodného algoritmu tyto komponenty rozdělit na dvě samostatné skupiny. Komponenty, které mají větší energii na levém mikrofónu, lze seřadit do skupiny komponent, které přišly zleva, a ty, co mají větší energii na pravém mikrofónu, přiřadit skupině komponent, které přišly zprava.

Výše uvedeným postupem je tedy uplatněna metoda slepé separace zdrojů k určení směru, ze kterého přišel zdroj (komponenta) k mikrofónu. V dalším kroku se této směrové informace o jednotlivých komponentách využije k efektivnější rekonstrukci původních signálů a tedy k vylepšení samotné metody slepé separace zdrojů.



### 3.1.1 Původní provedení metody BSS

V této kapitole je popsána metoda slepé separace, ze které vychází navrhované řešení uvedené v následujících dvou kapitolách a která byla v základech představena v teoretické části této práce (konkrétně v kapitole 2.2). Je zde umístěna také pro možnost srovnání různých řešení problému třídění a shlukování komponent.

Na začátku je vytvořena datová matice. Její řádky jsou sestaveny ze zaznamenaných vzorků signálů pomocí mikrofونů. Tyto vzorky mají zpoždění (časové posunutí) vždy o jeden vzorek, vyjádřené vztahem (4) uvedeným v kapitole 2.2. Celkové rozměry datové matice jsou dány počtem mikrofونů, délkou filtru a množstvím vzorků. Tato matice umožňuje najít vhodné filtry k získání nezávislých komponent ze zaznamenaných vzorků původních signálů pomocí ICA metody v dalším kroku.

Následně je tedy na datovou matici aplikována některá z metod pro analýzu nezávislých komponent (ICA). Tyto metody jsou blíže popsány v kapitole 2.2.1. Díky ICA je z datové matice získána demixovací matice, jejíž rozměry (počet sloupců a řádků) jsou dány počtem řádků datové matice. Vzájemným vynásobením této demixovací matice a původní datové matice vznikne matice nezávislých komponent. Tento krok využívá počáteční předpoklad ICA metody, a sice že původní signály jsou nezávislé.

Na získanou matici nezávislých komponent je aplikován některý z klastrovacích algoritmů. Tento krok je nazýván shlukování komponent. Konkrétně tato metoda shluje komponenty na základě jejich podobnosti (bližší popis v kapitole 2.2.2) narozdíl od navrhované metody, která využívá informace o směrovosti komponent. Účelem tohoto kroku je vypočítat matici podobnosti jednotlivých nezávislých komponent a zjistit tedy, které komponenty patří ke stejnému původnímu zdroji a ty následně seskupit do shluku, aby mohly být v dalším kroku pomocí nich zrekonstruovány původní zdroje. Každý takový shluk nezávislých komponent potom představuje jeden původní zdroj. Počet shluků by tedy měl být roven počtu původních zdrojů.

Dalším krokem je zpětná rekonstrukce. Kvůli neznámé filtraci nezávislých komponent je nemožné jejich použití pro odhadnutí původních signálů. Rekonstruuji se

proto odezvy od jednotlivých původních zdrojů na jednotlivých mikrofonech. Tyto odezvy jsou definovány vztahem (2). Jejich rekonstrukce se provádí z komponent seskupených ve shluku, získaném v minulém kroku. Jak bylo řečeno výše, jeden shluk představuje jeden zdroj. Je tedy nutné tento krok provést pro každý shluk komponent zvlášť.

Postupuje se tak, že řádky matice nezávislých komponent, získané v kroku po aplikaci ICA metody, se vynásobí vahou vyjadřující příslušnost komponent k danému shluku a takto upravená matice se následně násobí inverzní demixovací maticí, vytvořenou dříve pomocí ICA algoritmu (váhování viz. kapitola 2.2.3). Výstupem je rekonstruovaná původní datová matice, která byla vytvořena na samém začátku ze zaznamenaných vzorků signálů, ale pouze z komponent, které patřily k danému shluku. V případě dvou zdrojů tedy tyto zrekonstruované matice vzniknou dvě (8). Odhad konkrétní odezvy na určitém mikrofону je poté určen jako (9).

Na závěr se do jednoho audio kanálu sečtou odezvy jednotlivých původních zdrojů z každého mikrofónu, aby se dosáhlo co největší kvality původního signálu a maximálně se potlačilo rušení okolního prostředí.

### **3.1.2 Využití metody BSS k odhadu směru komponent**

Počáteční postup metody je shodný s metodou představenou v kapitole 3.1.1. Je tedy vytvořena datová matice zaznamenaných vzorků signálů, na kterou je následně aplikována některá z ICA metod a tak získána demixovací matice. Další kroky se však už od výše zmíněné metody liší.

Následujícím krokem je odhad směru, ze kterého přišly jednotlivé komponenty původních zdrojů. Tento krok je tedy nutné provést pro každou komponentu zvlášť. Celkový počet provedení je roven počtu komponent, tedy součinu počtu mikrofónů a délky separujícího filtru. Postupuje se tak, že se vždy pouze té jedné právě počítané komponentě přiřadí váha o hodnotě jedna a všem ostatním nula. V podstatě je tedy

tímto způsobem každá komponenta označena jako shluk a rekonstrukce odezvy na mikrofonu bude počítána pouze z této jedné komponenty.

Nejprve dojde k rekonstrukci původní datové matice pomocí shluku, ve kterém je pouze tato jedna komponenta a to tak, že inverzní demixovací matice se vynásobí vahou příslušnosti komponent ke shluku a výsledek se násobí původní demixovací maticí. Výstupem je zrekonstruovaná datová matice, avšak pouze z této jedné komponenty. Následně je spočten odhad odezev na jednotlivých mikrofonech pro tuto jednu komponentu.

Z těchto odezev, konkrétně z jejich energie, je určeno, ze kterého směru komponenta přišla k mikrofonom. Komponenta má totiž odlišnou energii na různých mikrofonech a podle poměru těchto energií lze odhadnout její směrovost. Ještě je nutné provést normalizaci těchto hodnot, aby nedocházelo ke zkreslení, pokud například jeden z mikrofónů bude mít nastavené vyšší předzesílení než druhý.

Normalizace se provede tak, že se sečtou hodnoty energie všech komponent na prvním a na druhém mikrofónu. Následně se hodnota energie každé komponenty na prvním i druhém mikrofónu vydělí příslušným součtem. Každá komponenta má nyní vyjádřen směr dvěma hodnotami, energií na levém mikrofónu a energií na pravém mikrofónu.

	Energie na levém mikrofónu	Energie na pravém mikrofónu
Komponenty	0.0011	0.0003
	0.0012	0.0001
	0.0000	0.0007
	0.0001	0.0030
	0.0001	0.0038
	0.0012	0.0001
	0.0013	0.0002
	0.0017	0.0001
	0.0012	0.0001
	0.0024	0.0002
	0.0022	0.0002
	0.0136	0.0034
	0.0003	0.0059
	0.0002	0.0029
	0.0002	0.0049
	0.0002	0.0029
	0.0162	0.0080
	0.0199	0.0144
	0.0010	0.0054
	0.0003	0.0062

Obrázek 5 Příklad matice znázorňující energii komponent na mikrofonech

Zdroj: vlastní zpracování

Pro další postup je ale nutné vyjádřit směr komponenty pouze jedním číslem. Byl zvolen interval čísel (0; 1). Komponenty s hodnotou směrovosti v intervalu (0; 0,5) jsou ty, které přišly zleva, komponenty ležící v intervalu (0,5; 1) jsou určeny jako směřované zprava. Tato hodnota směrovosti komponenty je dána vztahem

$$\frac{\frac{B}{A}}{\frac{B}{A} + 1}, \quad (11)$$

kde B je hodnota energie na pravém mikrofonu a A je hodnota energie na levém mikrofonu.

Tímto způsobem je určena směrovost všech komponent a o každé komponentě je tedy známo, zda patří ke zdroji, jehož signál dorazil na mikrofony zleva, nebo ke zdroji, jehož signál přišel zprava. Této znalosti směrovosti jednotlivých komponent se využije v dalším kroku, kterým je rekonstrukce odezev původních zdrojů.

### 3.1.3 Vylepšení metody BSS s využitím odhadu směru zdrojů

Závěrečným krokem je rekonstrukce odezev původních zdrojů na mikrofonech. K této rekonstrukci budou využity jednotlivé komponenty a informace o jejich směru tak, jak byly získány v minulém kroku.

Index směrnosti	
Komponenty	0.2112
	0.1021
	0.9874
	0.9695
	0.9822
	0.0523
	0.1556
	0.0668
	0.0569
	0.0739
	0.0836
	0.1995
	0.9543
	0.9511
	0.9674
	0.9398
	0.3299
	0.4206
	0.8466
	0.9498

Obrázek 6 Příklad matice znázorňující směrnost komponent

Zdroj: vlastní zpracování

Samotná rekonstrukce odezev původních zdrojů proběhne celkem dvakrát. Jednou pro každý původní zdroj. Konkrétně například pro pravý zdroj se do jednoho shluku seskupí všechny komponenty, které mají hodnotu směrnosti větší než 0,5. Všem komponentám v tomto shluku je následně přiřazena váha hodnoty jedna a ostatním komponentám váha hodnoty nula. Následně je spočtena rekonstrukce původní datové matice a odhad odezev tohoto zdroje na jednotlivých mikrofonech, a to stejným způsobem, jakým to bylo provedeno o několik kroků zpět pro jednotlivé komponenty.

Závěr je již shodný s metodou popsanou v kapitole 3.1.1 a jedná se tedy o součet jednotlivých odezev původních zdrojů na mikrofonech do jednoho audio kanálu.

## ***3.2 Implementace řešení v prostředí MATLAB***

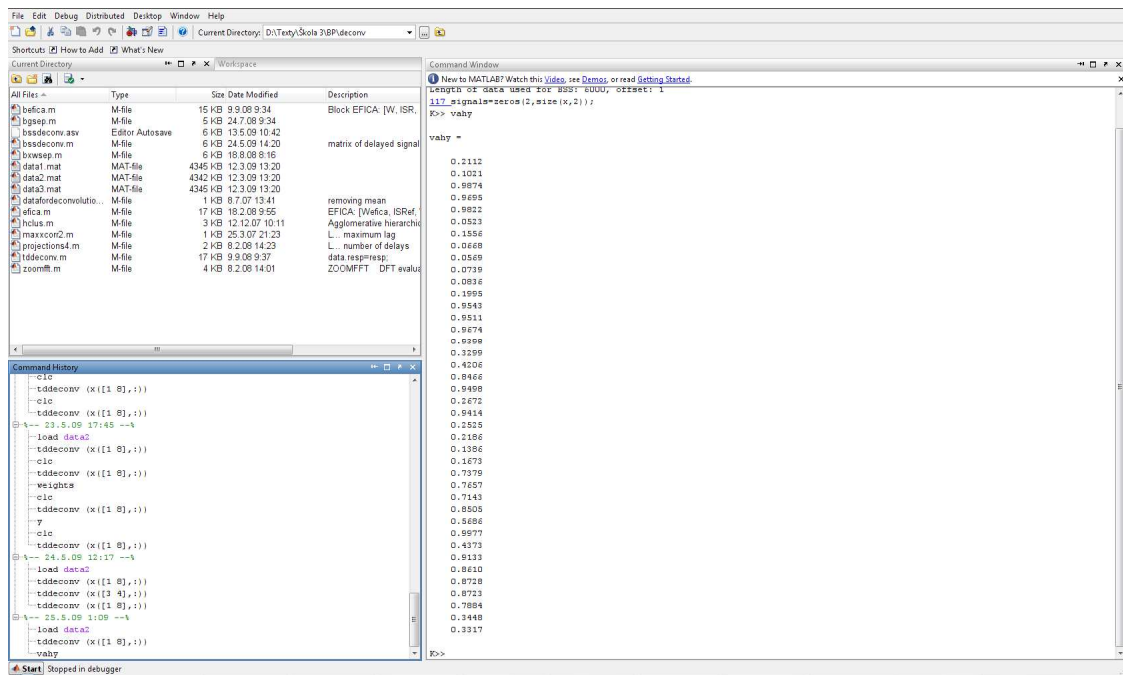
V této kapitole je pro názornost na obrázcích z prostředí aplikace MATLAB popsána implementace algoritmu navrženého pro řešení zadaných cílů. Tento algoritmus vznikl upravením kódu určeného pro původní provedení metody BSS, jež nevyužívala znalosti směru komponent, ale informaci o jejich vzájemné podobnosti.

### **3.2.1 Vývojové prostředí MATLAB**

MATLAB (MATrix LABoratory) je vyšší programovací jazyk pro technické výpočty a interaktivní prostředí určené k vývoji algoritmů, grafické vizualizaci, analýze dat a numerickým výpočtům. S použitím tohoto vývojového prostředí je možné řešit technické výpočetní problémy rychleji než s tradičními programovacími jazyky, jakou jsou C, C++ a Fortran.

Nachází široké uplatnění zejména v oblasti zpracování signálů a obrazu, navrhování komunikačních a řídicích systémů, testování a měření, finančního modelování, finanční analýzy a výpočetní biologie. Speciální doplňkové funkce, rozšiřující prostředí MATLAB, umožňují řešit specifické druhy problémů ve výše uvedených oblastech použití. Užitečná je také možnost návrhu vlastního grafického rozhraní pro vytvořené aplikace.

Je v něm obsažena celá řada prvků pro dokumentaci a sdílení práce uživatelů. Kód je možné sjednotit s dalšími programovacími jazyky a aplikacemi, jako jsou například C, C++, Fortran, Java nebo Microsoft Excel, a vzájemně tak přenášet jednotlivé algoritmy. [4]



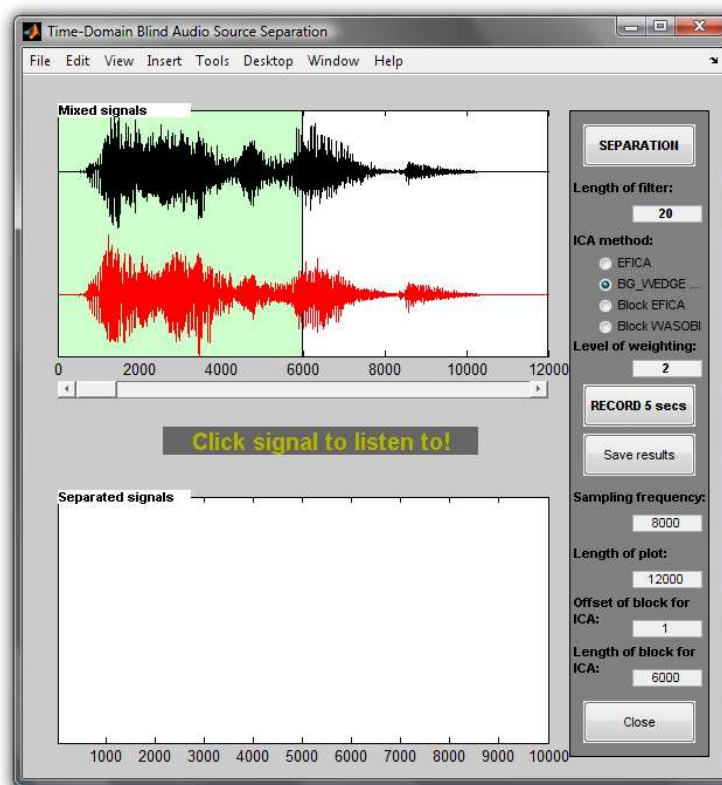
*Obrázek 7 Rozhraní aplikace MATLAB*

Zdroj: MATLAB

MATLAB byl vytvořen v pozdních sedmdesátých letech Clevem Molerem, vedoucím katedry informatiky Univerzity New Mexico. Vynález se brzy rozšířil na další univerzity a získal si silné publikum v komunitě lidí zabývajících se aplikovanou matematikou. Komerční potenciál tohoto programovacího jazyka využil až inženýr Jack Little, který v roce 1984 společně s Clevem Molerem a Stevem Bangertem založil společnost The MathWorks.

Jako první si MATLAB osvojili inženýři vyvíjející řídicí systémy, ale postupně se začal uplatňovat i v dalších oblastech. V současnosti je využíván i ve školství, zejména ve vyučování lineární algebry a numerické analýzy, a je populární mezi vědci zabývajícími se zpracováním obrazu. [5]

### 3.2.2 Vzhled aplikace a zdrojový kód



Obrázek 8 Dialogové okno aplikace pro BSS z prostředí MATLAB

Zdroj: MATLAB

Na výše uvedeném obrázku je zobrazeno dialogové okno, které po spuštění aplikace nabídne několik voleb pro upřesnění požadované separace. Je možné nastavit délku filtru, zvolit ICA metodu, nahrát 5 sekund dlouhý audio záznam, uložit výsledek separace, nastavit vzorkovací frekvenci, počet vzorků pro separaci nebo délku úseku zobrazených signálů. Kliknutím na některý z vyobrazených signálů je možné si ho přehrát a poslechnout. Lze si pustit jak původní smíšené signály, tak i jednotlivé signály po provedení separace. Volba úrovně váhování komponent se vztahuje k původní metodě slepé separace, která využívala podobnosti komponent. Pro navrženou metodu pracující na základě informace o směru komponent již tato volba nemá význam.

Dále následuje popis jednotlivých kroků navrženého algoritmu, tak jak postupně probíhají po stisknutí tlačítka „SEPARATION“:



```

1 function [signals, microphones]=bssdeconv(x,L,LengthOfICA,offset,method,wlevel,hwait)
2 fprintf('Length of data used for BSS: %d, offset: %d\n',LengthOfICA,offset);
3
4 y=datafordeconvolution(x(:,offset:offset-1+LengthOfICA),L,1);
5
6 waitbar(0.3,hwait,'ICA decomposition ...');
7 if strcmp(method,'bgl')
8     We=bgssep(y,min([floor(length(y)/300) 40]));
9 elseif strcmp(method,'efica')
10    We=efica(y,eye(size(y,1)));
11 elseif strcmp(method,'extefica')
12    We=befica(y,floor(length(y)/300),eye(size(y,1)));
13 elseif strcmp(method,'bwasobi')
14    We=bxwsep(y,floor(length(y)/300),1);
15 end
16 A=inv(We);
17

```

Obrázek 9 Řádky 1 – 17 zdrojového kódu algoritmu v prostředí MATLAB

Zdroj: MATLAB

- řádky 1 - 2: načtení parametrů pro separaci nastavených v aplikaci
- řádek 4: vytvoření datové matice z vybraného úseku signálů
- řádky 6 - 15: aplikace zvolené ICA metody a vytvoření demixovací matice
- řádek 16: vytvoření inverzní demixovací matice

```

18 waitbar(0.6,hwait,'Estimating direction of components ...');
19 for i=1:(size(x,1)*(L+1))
20
21     weights=zeros((size(x,1)*(L+1)),1);
22     weights(i)=1;
23
24     H=(A*diag(weights))*We;
25
26     for el=1:size(x,1)
27         for j=1:size(x,1)
28             for k=1:L+1
29                 HH(k,:)=conv(H((el-1)*(L+1)+k,(j-1)*(L+1)+1:j*(L+1)),[zeros(1,L-k+1) 1 zeros(1,k-1)]);
30             end
31             h(el,:,j)=sum(HH);
32         end
33     end
34
35     microphones=zeros(length(x),size(x,1));
36
37     for el=1:size(x,1)
38         for j=1:size(x,1)
39             microphones(:,el)=microphones(:,el)+filter(h(el,:,j),1,x(j,:));
40         end
41     end
42

```

Obrázek 10 Řádky 18 – 42 zdrojového kódu algoritmu v prostředí MATLAB

Zdroj: MATLAB

- řádek 19: začátek cyklu pro výpočet energie komponenty na mikrofonu
- řádky 21 - 22: označení komponenty jako shluk
- řádky 24 - 33: výpočet transformace pro rekonstrukci datové matice
- řádky 35 - 41: odhad odezev na mikrofonech

```

43 - a=squeeze(microphones);
44 - polohy(i,:)=mean(a.^2);
45 - end
46
47 - b=sum(polohy);
48 - for i=1:(size(x,1)*(L+1))
49 -     polohy2(i,1)=polohy(i,1)/b(1,1);
50 -     polohy2(i,2)=polohy(i,2)/b(1,2);
51
52 -     vahy(i,:)=(polohy2(i,2)/polohy2(i,1))/((polohy2(i,2)/polohy2(i,1))+1);
53 - end
54

```

Obrázek 11 Řádky 43 – 54 zdrojového kódu algoritmu v prostředí MATLAB

Zdroj: MATLAB

- řádek 43: převod signálu na 2-D maticový tvar tak, aby v každém sloupci byly hodnoty pro jeden mikrofon
- řádek 44: výpočet energie komponenty na mikrofonech
- řádek 45: konec cyklu pro výpočet energie komponenty na mikrofonu
- řádek 47: součet energií komponent na mikrofonech
- řádky 48 - 50: normalizace hodnot energie komponenty
- řádek 52: výpočet indexu směrovosti komponenty

```

55 - signals=zeros(2,size(x,2));
56 - microphones=zeros(2,length(x),size(x,1));
57
58 - waitbar(0.8,hwait,'Reconstruction ...');
59
60 - for i=1:2
61
62 -     if i==1
63 -         incluster=find(vahy>0.5);
64 -     else
65 -         incluster=find(vahy<0.5);
66 -     end
67 -     outcluster=setdiff(1:(size(x,1)*(L+1)),incluster);
68
69 -     weights(incluster)=1;
70 -     weights(outcluster)=0;
71

```

Obrázek 12 Řádky 55 – 71 zdrojového kódu algoritmu v prostředí MATLAB

Zdroj: MATLAB

- řádky 55 - 56: reset signálů a odezev na mikrofonech
- řádek 60: začátek cyklu pro rekonstrukci zdroje
- řádky 62 - 67: rozdělení komponent podle směru
- řádky 69 - 70: váhování komponent

```

72 - H=(A*diag(weights))*We;
73
74 - for el=1:size(x,1)
75 -     for j=1:size(x,1)
76 -         for k=1:L+1
77 -             HH(k,:)=conv(H((el-1)*(L+1)+k,(j-1)*(L+1)+1:j*(L+1)),[zeros(1,L-k+1) 1 zeros(1,k-1)]);
78 -         end
79 -         h(el,:,j)=sum(HH);
80 -     end
81 - end
82
83 - for el=1:size(x,1)
84 -     for j=1:size(x,1)
85 -         microphones(i,:,el)=microphones(i,:,el)+filter(h(el,:,j),1,x(j,:));
86 -     end
87 - end
88

```

Obrázek 13 Řádky 72 – 88 zdrojového kódu algoritmu v prostředí MATLAB

Zdroj: MATLAB

- řádek 72 - 81: výpočet transformace pro rekonstrukci datové matice
- řádky 83 - 87: odhad odezev na mikrofonech

```

89 - [val lag]=maxxcorr2(squeeze(microphones(i,:,:)'),L);
90 - formmics=microphones(i,:,1);
91 - for k=2:size(x,1)
92 -     if lag(1,k)>0
93 -         formmics(lag(1,k)+1:end)=formmics(lag(1,k)+1:end)+microphones(i,1:end-lag(1,k),i);
94 -     else
95 -         formmics(1:end+lag(1,k))=formmics(1:end+lag(1,k))+microphones(i,lag(k,1)+1:end,k);
96 -     end
97 - end
98 - formmics=formmics/size(x,1);
99 - signals(i,:)=formmics;
100 - end

```

Obrázek 14 Řádky 89 – 100 zdrojového kódu algoritmu v prostředí MATLAB

Zdroj: MATLAB

- řádky 89 - 99: součet odezev do jednoho signálu
- řádek 100: konec cyklu pro rekonstrukci zdroje

## 4 Experimenty

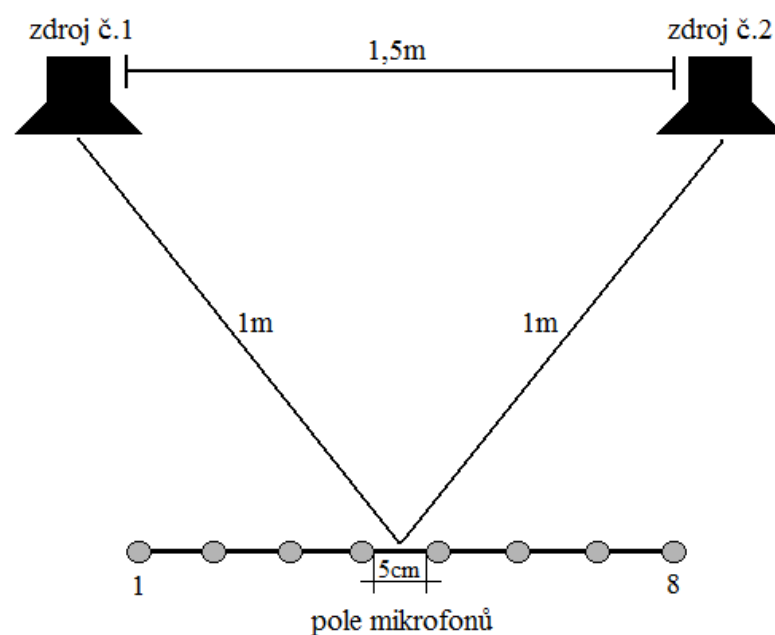
Tato kapitola se zabývá testováním navržené metody slepé separace na různých audio nahrávkách. V první části kapitoly je vysvětleno, v jakých podmínkách byly testované audio nahrávky pořízeny a pomocí kterých kritérií jsou jejich separace hodnoceny. Druhá část kapitoly obsahuje přehled dosažených výsledků a jejich vyhodnocení.

### *4.1 Vlastnosti a kritéria provedených experimentů*

Směsi audio nahrávek byly pořízeny s použitím dvou reproduktorů a lineárního pole mikrofónů, na které se signály zaznamenávaly. Toto pole tvoří celkem osm mikrofónů, aby bylo možné volbou různých dvojic měnit podmínky pro separaci signálů. Orientační podoba této aparatury je vyobrazena na obrázku 15.

Pro zhodnocení dosažené kvality provedené separace slouží dvě kritéria a to SIR (signal-to-interference) a SDR (signal-to-distortion). Hodnota SIR vyjadřuje poměr mezi signálem a rušením. V tomto případě jsou v signálu rušivým elementem zbytky druhého signálu, které v něm zůstaly po separaci. Parametr SDR vyjadřuje zkreslení signálu po separaci oproti původnímu. Je tím větší, čím menší je vzdálenost separovaného signálu od chtěného. Hodnoty SIR a SDR jsou udávány v decibelech a pro obě platí čím větší, tím lepší je kvalita separovaných signálů.

Vyhodnocení separace a tedy zjištění hodnot SIR a SDR bylo provedeno pomocí nástroje BSS evaluation toolbox [7] v prostředí programu MATLAB, který k tomuto účelu využívá původní nesmíchané signály (zdrojové soubory a bližší informace o tomto nástroji na internetové adrese <http://www.irisa.fr/metiss/gribonval/b>).

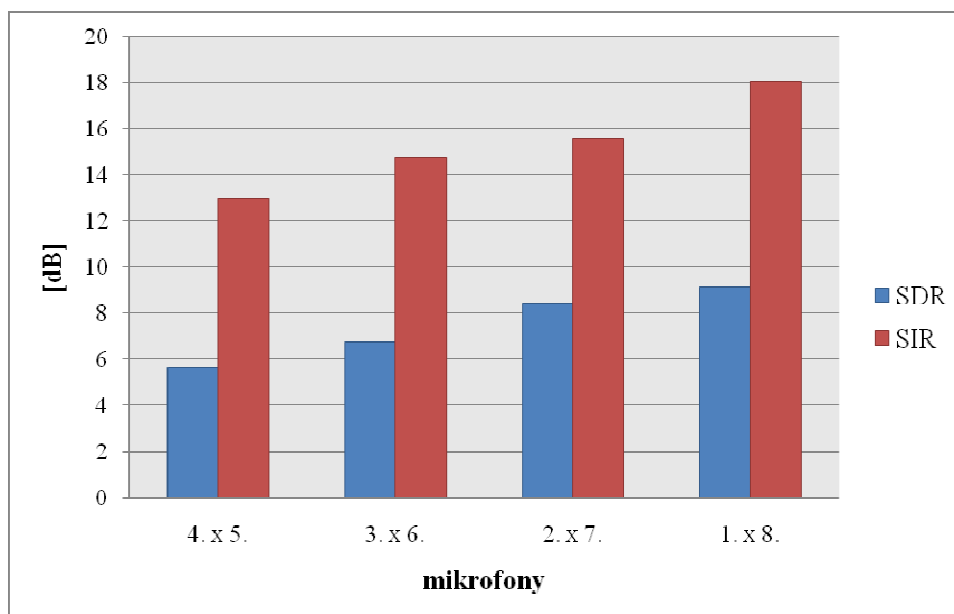


Obrázek 15 Pole mikrofonů pro pořizování smíšených audio záznamů

Zdroj: vlastní zpracování

## 4.2 Vyhodnocení provedených separací

První experiment byl proveden na smíšeném záznamu hovořícího muže a rušivého zvuku. Účelem bylo zjistit, jak ovlivňuje kvalitu separace vzdálenost mikrofonů, pomocí kterých byl záznam pořízen. Separace tedy probíhala vždy při stejném nastavení, měnila se pouze dvojice mikrofonů, pomocí kterých byl nahrán záznam směsi. Výsledné hodnoty jsou uvedeny v grafu 1.

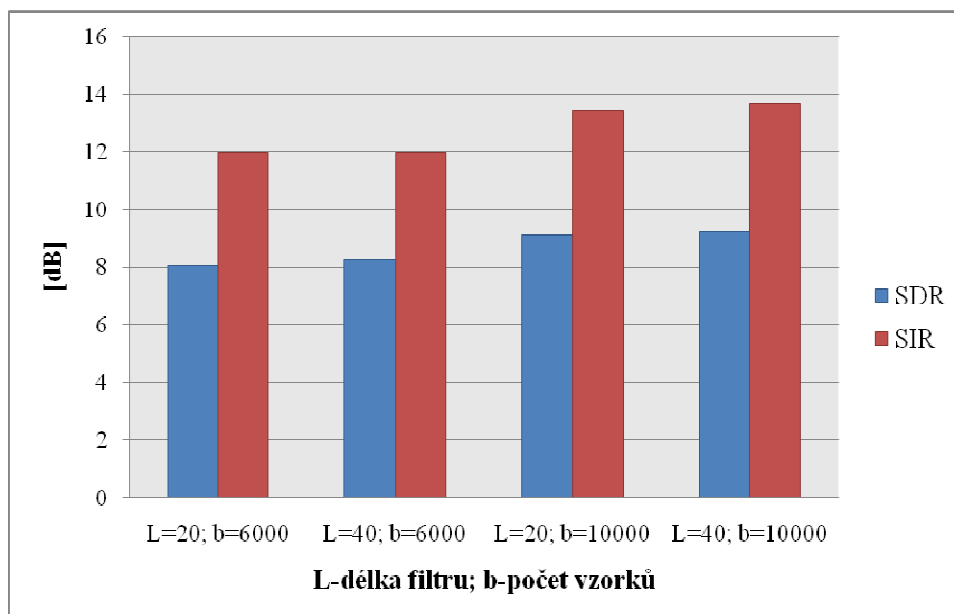


*Graf 1 Hodnoty SIR a SDR v závislosti na použité dvojici mikrofonů*

Zdroj: vlastní zpracování

Z hodnot na výše vyobrazeném grafu je vidět, že navržená metoda pracuje tím lépe, čím větší je vzdálenost mezi dvojicí použitých mikrofonů. Tento jev je způsoben snadnějším rozdělením komponent původních zdrojů do shluků následně použitých k rekonstrukci, protože na vzdálenějších mikrofonech má každá komponenta výraznější poměr mezi energií na jednom a na druhém mikrofону. Výsledná separace je tak přesnější a signál neobsahuje tolik rušivých elementů, tedy zbytků druhého signálu, jako v případě navzájem si blízkých mikrofonů, kdy je toto rušení poměrně vysoké.

K provedení druhého experimentu byl použit smíšený záznam dvou různých zvuků. Tentokrát bylo cílem zjistit, jak moc navržená metoda reaguje na změny v nastavení separace. Záznam byl tedy separován vždy z nahrávky pořízené stejnými mikrofóny, konkrétně třetím a šestým, měnila se délka filtru a počet vzorků použitých pro separaci. Závislost výsledných hodnot na těchto změnách ukazuje graf 2.



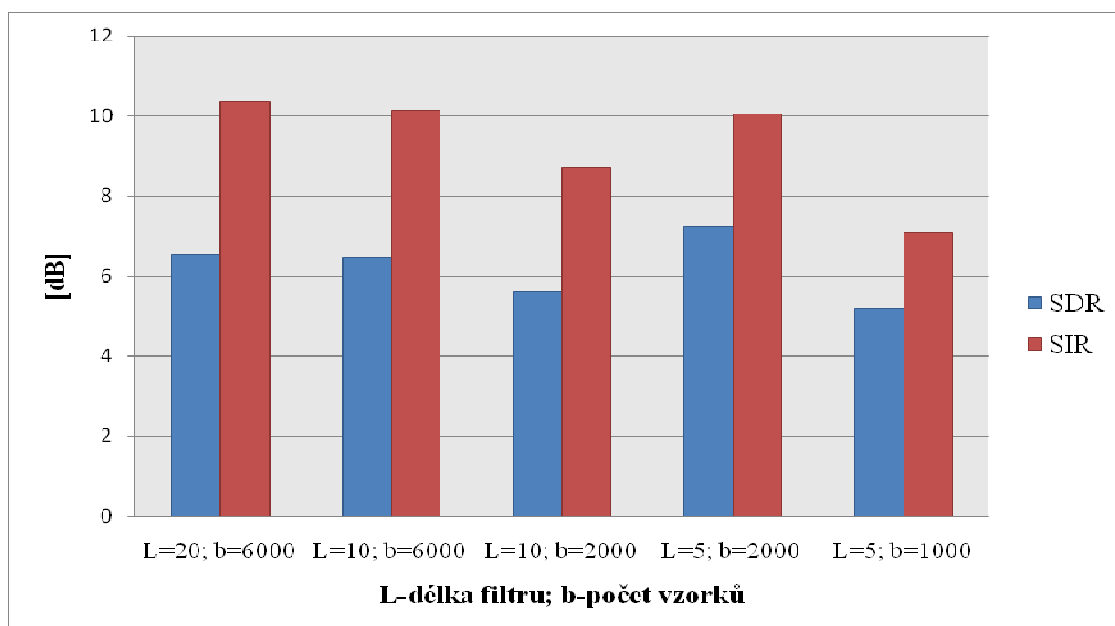
Graf 2 Hodnoty SIR a SDR v závislosti na použitém nastavení

Zdroj: vlastní zpracování

Ukázalo se, že v případě tohoto experimentu, má délka filtru a s ní související počet komponent použitých pro rekonstrukci odezev původních zdrojů na mikrofonech na výsledné hodnoty velmi malý dopad. Hodnoty separace s delším filtrem vykazují pouze nepatrné zlepšení. Zvýšení počtu komponent by mělo větší vliv na výsledky u původní metody s rekonstrukcí založenou na vzájemné podobnosti mezi komponentami, ale u navrženého řešení pracujícího na základě směrovosti komponent nemá navýšení počtu komponent příliš význam. Jejich počet neovlivní uspořádání shluků pro rekonstrukci. K určitému zlepšení dojde při použití většího počtu vzorků neboli delšího úseku záznamu. Díky tomuto navýšení je zvýšena efektivita použité ICA metody. S navýšením délky datové matice ovšem roste i čas potřebný na zpracování.

První dva experimenty probíhaly za relativně příznivých podmínek. Délka filtru neklesla pod hodnotu 20 a délka záznamu nebyla menší než 6000 vzorků což je při použité vzorkovací frekvenci 8000Hz přibližně jedna sekunda. Následující experiment měl tedy za cíl zjistit, jak si navržená metoda poradí s velmi krátkým filtrem a malým počtem vzorků zaznamenaných signálů. Nahrávka určená k tomuto experimentu obsahovala smíšený záznam hovořícího muže a hrající hudby. Použity byly mikrofony číslo 1 a 8. Výsledky jsou obsaženy v grafu 3.





Graf 3 Hodnoty SIR a SDR v závislosti na zkrácení délky separovaného signálu

Zdroj: vlastní zpracování

Hodnota délky filtru se zpočátku na výsledcích příliš neprojevovala, stejně jako tomu bylo v minulém experimentu. Se snižující se délkou záznamu už byl pokles kvality separace výraznější, nicméně nijak dramatický. Zajímavý moment ovšem nastal při čtvrtém měření, kdy byl filtr nastaven na délku 5 a počet vzorků záznamu byl roven hodnotě 2000. Výsledná separace se hodnotou SIR vyrovnala pokusu s dvojnásobně dlouhým filtrem a třikrát větším počtem vzorků. Hodnotou SDR dokonce tento pokus překonala.

Tento jev byl pravděpodobně způsoben tím, že konkrétně u tohoto záznamu měly komponenty použité pro rekonstrukci velmi výrazný poměr energií na mikrofonech a kvalita výsledné separace byla tedy vyšší než u měření předchozích, které sice měly k dispozici více komponent, ovšem s horším určením jejich směrovosti. Jistě tomuto výsledku také napomohly použité mikrofony, konkrétně 1. a 8., které jsou od sebe dostatečně vzdálené pro kvalitní určení směru komponent. Při použití jiných, navzájem si bližších mikrofonů by výsledky nejspíše dopadly daleko hůře.

Jak je vidět z výsledků měření, kvalita separace je při použití vzdálených mikrofonů dostačující i s nastavením velmi krátkého filtru a malé délky záznamu. Nároky na výpočetní výkon jsou velmi malé a separace je tak provedena velice rychle.

## 5 Závěr

Tato práce se zabývala problémem slepé separace akustických zdrojů ze smíšených záznamů pořízených mikrofony. Oblast zpracování digitálních signálů je velmi obsáhlá, proto bylo nutné v první části práce popsat nezbytnou teorii potřebnou k porozumění problému.

Dále byla představena jedna z funkčních metod řešení problému, která je založena na vzájemné podobnosti jednotlivých komponent akustických signálů. Z této metody vychází navržené řešení, které je postaveno na myšlence využití směrovosti komponent. Jedná se pouze o jednu z mnoha možných cest řešení tohoto problému.

Výsledkem je funkční algoritmus, navržený konkrétně pro separaci dvou zdrojů pomocí dvou mikrofónů. Efektivita tohoto algoritmu byla ověřena provedením několika experimentálních separací. Navržené řešení nabízí jeden z možných pohledů na problém slepé separace a je možné ho dále upravovat a vylepšovat. Může posloužit i ostatním kolegům zabývajícím se touto problematikou jako experimentální pomůcka nebo důvod k zamyšlení se nad problémem z jiného pohledu.

Problém slepé separace je i nadále oblastí ze značné míry neprobádanou a jeho pokračující výzkum může přinést velmi užitečné vynálezy v oblasti audio techniky, pomůcek pro handicapované nebo nové přístroje, které najdou využití v medicíně.

## Seznam použité literatury

[1] EKSLER, V. *Analýza hlavních komponent v problematice separace naslepo* [online]. [cit. 16.5.2009].

URL: <<http://www.elektrorevue.cz/clanky/05029/index.html>>.

[2] HYVÄRINEN, A. *What Is Independent Component Analysis?* [online].

[cit. 16.5.2009]. URL: <<http://www.cs.helsinki.fi/u/ahyvarin/whatisica.shtml>>.

[3] KOLDOVSKÝ, Z., TICHAVSKÝ, P. *Time-domain Blind Audio Source Separation Using Advanced Component Clustering and Reconstruction*. Proc. of The Joint Workshop on Hands-free Speech Communication and Microphone Arrays (HSCMA 2008), pp. 216 – 219. Trento, Italy, 2008. ISBN 978-1-4244-2338-5.

[4] MATHWORKS. *MATLAB – Introduction and Key Features* [online].

[cit. 16.5.2009].

URL: <<http://www.mathworks.com/products/matlab/description1.html>>.

[5] MATHWORKS. *MATLAB News & Notes – December 2004 – Cleve's Corner – The Origins of MATLAB* [online]. [cit. 16.5.2009].

URL: <[http://www.mathworks.com/company/newsletters/news\\_notes/clevescorner/dec04.html](http://www.mathworks.com/company/newsletters/news_notes/clevescorner/dec04.html)>.

[6] PORAT, B. *A Course In Digital Signal Processing*. New York, John Wiley & Sons, 1997. ISBN 0471149616.

[7] VINCENT, E., FÉVOTTE, C. and GRIBONVAL, R. *Performance measurement in Blind Audio Source Separation*. IEEE Trans. Audio, Speech and Language Processing, 14(4), pp. 1462-1469, 2006.

[8] WIKIPEDIA. *Colin Cherry* [online]. [cit. 16.5.2009].

URL: <[http://en.wikipedia.org/wiki/Colin\\_Cherry](http://en.wikipedia.org/wiki/Colin_Cherry)>.

[9] WIKIPEDIA. *Computational auditory scene analysis* [online]. [cit. 16.5.2009].

URL: <[http://en.wikipedia.org/wiki/Computational\\_auditory\\_scene\\_analysis](http://en.wikipedia.org/wiki/Computational_auditory_scene_analysis)>.

[10] WIKIPEDIA. *Independent component analysis* [online]. [cit. 16.5.2009].

URL: <[http://en.wikipedia.org/wiki/Independent\\_components\\_analysis](http://en.wikipedia.org/wiki/Independent_components_analysis)>.

[11] WIKIPEDIA. *Magnetoencephalography* [online]. [cit. 16.5.2009].

URL: <<http://en.wikipedia.org/wiki/Magnetoencephalography>>.

[12] WIKIPEDIA. *Principal component analysis* [online]. [cit. 16.5.2009].

URL: <[http://en.wikipedia.org/wiki/Principal\\_components\\_analysis](http://en.wikipedia.org/wiki/Principal_components_analysis)>.

[13] WIKIPEDIA. *Source separation* [online]. [cit. 16.5.2009].

URL: <[http://en.wikipedia.org/wiki/Source\\_separation](http://en.wikipedia.org/wiki/Source_separation)>.